Dear Reviewer:

We would like to express our gratitude to you for the helpful comments to improve this manuscript. We have carefully modified the discussion and the expression following your comments. The specific responses and revisions are shown below. They are in blue font for clarity.

Qinghua Yang, Qian Shi, Robert Ricker, Stefan Hendricks
On behalf of all the authors

GC1: Why not use same snow product and methods for ICESat and Envisat? Now the discussion of the differences due to sensors and snow depth only include a sensitivity of the hydrostatic equilibrium to snow depth/freeboard, but not the actual effect. It would be an option to calculate sea ice thickness from ICESat with the AMSR-E snow depths as well so you can compare what part of the difference is a direct effect from the difference in sensors and what is caused by the difference in snow depth. I understand that this involves quite some more work, but I think a the statement that is now made in the summary (L406-408) is a bit strong for the amount of proof you have for this, as you've not made the actual comparison.

Finally, we retrieve a new ICESat SIT using the same snow depth climatology product and the same hydrostatic equilibrium approach as Envisat SIT in response to GC1 suggested by RC3. The new ICESat SIT is compared with Envisat SIT and the variations of their differences are shown in Fig. 4. Compared to Fig. 4-6 in the manuscript, we can find that the positive differences in the Weddell Sea in summer and autumn increase, while the rest of the differences decrease.

However, we don't think this experiment is available to distinguish the impacts from the sensors and from snow depth as stated by RC3, since the new ICESat SIT still employs a possibly biased snow depth product. It can only be used to clarify the difference between the hydrostatic equilibrium retrieval method and the modified density retrieval method.

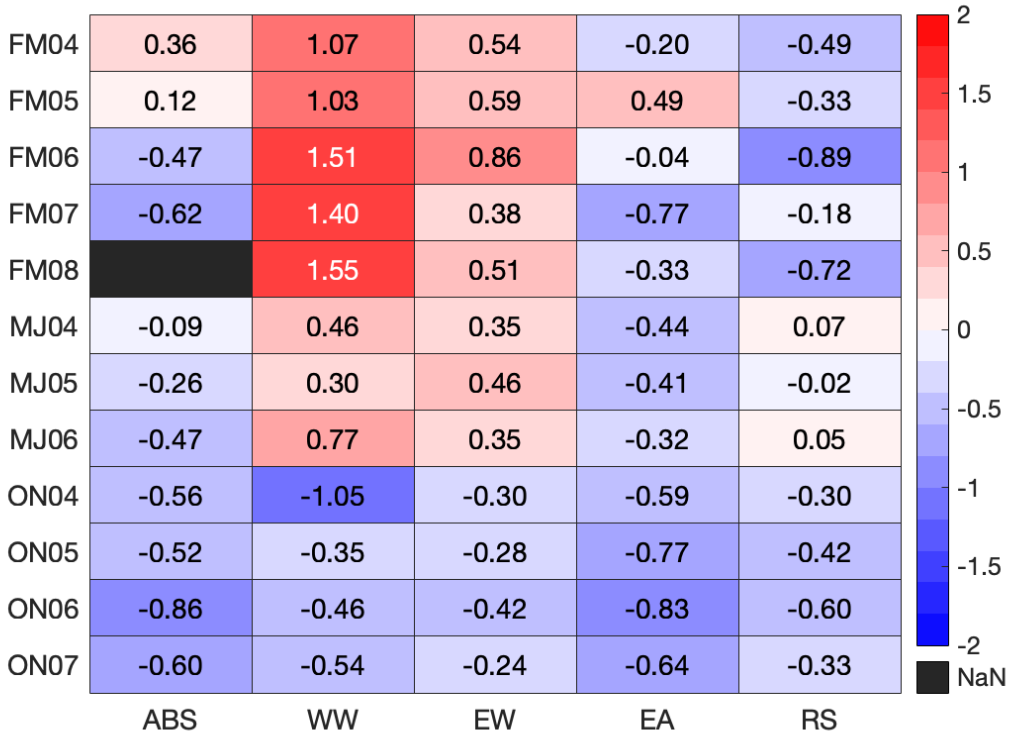


Fig. 1 Changes in the differences between Envisat and ICESat SIT for each comparison period and each region under the experiment of the new ICESat SIT.

GC2: Why is a different sea ice concentration threshold used for ICESat (60%, L143) than for Envisat (70%, L123)?

The usage of different SIC thresholds is because of the different thresholds used in the retrieval of the two data sets. Envisat SIT employs a SIC threshold of 70% during the retrieval while the ICESat SIT uses 60%. Only areas with sea ice concentrations greater than the threshold are considered a valid area for detection of leads and sea ice. We also tested the difference between using 60% and 70% SIC threshold for ICESat during the comparison with Envisat SIT. According to Table 1, this different threshold does not play an important role in the results of this paper. D(60) refers to Envisat minus ICESat (ENV-ICE) applying 60% SIC threshold for ICESat, while D(70) refers to ENV-ICE when SIC threshold for ICESat is 70%. Since the ice concentration gradients are usually quite steep, there will not be a lot of area with values 60% < SIC < 70%.

Table 1. Statistical results of the comparison between Envisat SIT and ICESat SIT using 60% and 70% SIC threshold at each operating period.

| | ON04 | ON05 | ON06 | ON07 | FM04 | FM05 | FM06 | MA07 | FM08 | MJ04 | MJ05 | MJ06 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| D(60) (m) | 0.00 | 0.05 | -0.19 | 0.14 | 0.89 | 0.74 | 0.47 | 0.61 | 0.92 | 0.61 | 0.55 | 0.60 |
| D(70) (m) | 0.00 | 0.06 | -0.21 | 0.15 | 0.79 | 0.66 | 0.43 | 0.61 | 0.89 | 0.60 | 0.55 | 0.61 |

GC3: In lines 381-386 you introduce an improvement of the method you have used to obtain ICESat sea

ice thickness. What is the reason for not using this improved method?

We investigated the ICESat product that Li et al. (2018) produces by comparing with ICESat from Kern et al. (2016) and the ULS SIT used in this study. From Fig. 2 we can see that the differences between two ICESat products are small in general, with some larger differences in the West Weddell Sea and Amundsen Sea. Table 2 shows that compared with ULS SIT, ICESat SIT from Kern et al. (2016) performs even better. Based on these analyses, we think that the inter-comparison results with Envisat SIT in this study are not affected by the choice of ICESat product. Therefore, our work is still based on the data produced by Kern et al. (2016).

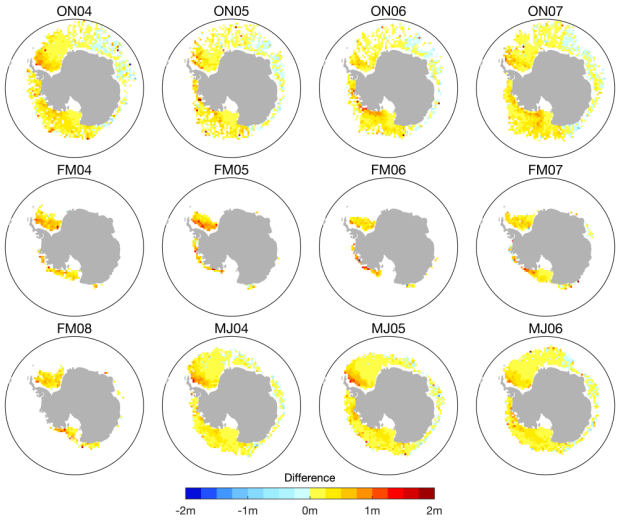


Fig. 2 Maps of differences that ICESat SIT from Kern et al. (2016) minus ICESat SIT from Li et al. (2018) at each operating period.

Table 2. The differences and RMSD between ULS SIT and the two ICESat SIT at each site. ICE(K) refers to ICESat SIT from Kern et al. (2016) and ICE(L) refers to ICESat SIT from Li et al. (2018).

| | ICE(K)-ULS | | ICE(L)-ULS | |
|---|---|---|---|---|
| | D (m) | RMSD (m) | D (m) | RMSD (m) |
| Site 207 | 0.60 | 0.66 | 0.68 | 0.34 |
| Site 229 | -0.04 | 0.07 | 0.17 | 0.15 |
| Site 231 | 0.27 | 0.34 | 0.33 | 0.44 |

GC4: L229: 'an overall comparison between Envisat and ICESat effective SIT'. In the methods it said the Envisat SIT product 'represents the actual SIT (i.e., mean thickness of the ice-covered fraction of the grid cell area)' (L122-123) and that the ICESat SIT product is the 'effective sea ice thickness (i.e., mean thickness per grid cell including open water ares)' (L141-142). Are these two products compared here? This would not be a fair comparison, as the effective sea ice thickness is by definition going to be thinner than the actual sea ice thickness. If you are comparing actual sea ice thickness products please clarify here and in the methods section.

During the comparison with ULS observations, we compare Envisat and ICESat actual SIT to exclude zero thickness measured by ULS. Therefore, we divide ICESat SIT by the sea ice concentration contained in the ICESat data for each grid.

Then, we compare the effective SIT of Envisat and ICESat during the intercomparison work by multiplying Envisat SIT by the sea ice concentration contained in the Envisat data. We added the

information in section 3.2: “The effective Envisat SIT is calculated by multiplying the sea ice concentration contained in the data for each grid which come from OSI-SAF Global Sea Ice Concentration (OSI-409) until April 16, 2015 and the OSI-SAF Global Sea Ice Concentration continuous reprocessing offline product (OSI-430) afterwards (http://osisaf.met.no).”

GC5: What values or products have been used for water, snow, and ice density in the calculations of sea ice thickness with Eq 1, Eq 2, and Eq 3? Are they the same for Envisat and ICESat? If not, discuss the effect on the results.
The $\rho$_water in Eq.1 is 1024 kg $m^{-3}$ while in Eq.3 is 1023.9 kg $m^{-3}$. The $\rho$_ice in Eq.1 is 916.7 kg $m^{-3}$ while in Eq. 2 is 915.1 kg $m^{-3}$. The $\rho$_snow in Eq.1 and Eq.2 are the same of 300 kg $m^{-3}$.

GC6: ULS and satellite altimetry SIT distributions would be interesting to see as well, if possible.
Thanks for your advice. However, since there are little corresponding ULS data, the large-scale SIT distributions cannot be presented.

GC7: There are some significant issues with interpreting sentences throughout the manuscript. I have added some key examples below, but the clarity of the manuscript could improve from a thorough read-through.
We checked through the manuscript and made our best to interpret every sentence clearly.

L77-78: 'Several freeboard- ... compared (Kern et al., 2016).' This sentence feels unrelated to the rest of the paragraph and is therefore confusing. If you want to go into this you need to explain the different retrieval algortihms. But I think it's better to leave this sentence out of the introduction and leave this to the methods (as you've explained this more clearly in section 2.2).
We agreed and removed this sentence as you suggested.

L148-149: 'The signals ... travel time.' This sentence makes it sound like travel time is used to differentiate observations of sea ice bottom vs. sea surface, but I think you are trying to say that the distance is determined from the travel time measurement. It also sounds like only two measurements are made, one from the sea ice bottom and one from the sea surface. Please rewrite this.
We rewrote the sentence: “The sensors transmit sound pulses upwards with a footprint of 6–8 m in diameter and the signals are reflected either by the sea ice bottom or the sea surface, **yielding two-way travel time which can be converted into distances.**”

L254: 'ICESat is more sensitive to thick ice than Envisat', but the Envisat SIT product is thicker than ICESat? You describe this bias well in section 4.1, but here it's a bit confusing, as you seem to say that ICESat should show thicker ice.
We realized that the statement is correct only for along-track data. Firstly, given the sparseness of ICESat overpasses with valid data such a 100 km grid SIT estimate in that region might be biased by the presence of thick landfast ice. Besides, ocean swell can result in anomalously high freeboard values which then convert into too high sea-ice thickness values. While this is a local phenomenon, the sparseness of ICESat overpasses with valid data can results in a similar effect as for landfast ice. Therefore, we considered the two issues here carefully and decided to remove this statement.

L262: 'Envisat has a positive difference with respect to ICESat'. I do not understand what this means. Have a look at the suggestions for technical corrections too.

We wrote the sentence: “In summer, the agreement between Envisat SIT and ICESat SIT is not good, mainly due to their different performances towards thick ice above 3 m.”

*Techinical corrections*

L24: 'while the uncertainties of *the* snow depth product are' or 'while the uncertainties of snow depth product*s* are'

We modified this sentence: “while the uncertainties of the snow depth product are not the dominant cause of the differences”.

L30: 'it is still unclear if ... sea ice thickness'. Change 'also associated with' to 'accompanied by', these changes do not have to be related (or associated) but can be seperate.

We modified this sentence: “However, it is still unclear if the recent increase in Antarctic sea ice area is also accompanied by a similar change in sea ice thickness.”

L39-40: 'from the ASPeCt can provide' change to 'by the ASPeCt expert group can provide'

We modified this sentence.

L42: 'airborne electromagnetic data which measure total freeboard', data don't measure things, maybe rephrase.

We changed “measure” to “provide”.

L52: Remove 'basically', this sounds very unscientific.

We removed this word.

L54-55: Consider more recent studies that have retrieved Antarctic sea ice thickness, e.g. Kurtz & Markus (2012) and Kacimi & Kwok (2020).

We added Kacimi & Kwok (2020) here and removed Zwally et al. (2008).

L83-84: 'also how the different ... distribution.' Very vague, what are 'the different retrieval methods', ICESat and Envisat?

Yes, the different retrieval methods that Envisat and ICESat SIT products use have impacts on their differences.

L88: 'the former inter-comparison study', which study is this?

The study is Kern et al. (2016), and we modified the sentence to make it clear: “Based on the former inter-comparison study (Kern et al., 2016), we choose the ICESat sea ice thickness data derived from the modified ice density approach for comparison.”

L147: change 'underwater' to 'below sea level'

We agreed with your comment but we found it unnecessary to mention the mooring location in detail.

L150-151: 'once several minutes', do you mean 'every several minutes'? Please rewrite and maybe be more specific (what is several minutes)?

We modified the sentence: “The intervals of sea ice draft measurements are between 3 and 15 minutes from November 1990 to March 2008.”

L153: seasons -> season

We corrected this issue.

L166: 'FM-MJ and MJ-ON'. I guess you are referring to February/March-May/June and May/June-October/November. Please specify the first time you mention these abbreviations.

We modified the sentence: “We compare FDD with the SIT variations from February/March to May/June (FM-MJ) and from May/June-October/November (MJ-ON) represented by Envisat and ICESat SIT.”

L182: 'Before ... first.' Repetitive, just use 'before' or 'first'.

We deleted “first”.

L193: Remove 'during the comparison'.

We removed this phrase.

L197: 'We provide ... SIT products.' Rewrite this sentence. I would suggest something like 'The error bars in the figure show the uncertainty estimates of/from the SIT products'.

We rewrote the sentence as: “The error bars in the figure show the uncertainty estimates from the SIT products.”

L197-200: 'The Envisat SIT ... Li et al., 2018).' Move these sentences to the methods? Also: I think adding an estimate of the ULS uncertainty to Figure 3 as well would improve the interpretation of this figure. You mentioned an estimate of the ULS uncertainty in

We added the ULS uncertainty of ± 0.05 m following Belter et al. (2020) in our paper.

L152-154. You now mention when the error bars of the altimetry sensors do not overlap with the ULS points, but it would be interesting to see if they do overlap with the ULS error bars.

We find that the ULS uncertainty cannot explain these differences.

L207: Why are the uncertainties of freeboard and snow depth not considered for the ICESat SIT uncertainties?

Snow depth uncertainty is not included because the ICESat SIT retrieval method does not require additional snow-depth information. However, we checked Kern et al. (2016) and their computation of the SIT uncertainty included in the product is based on uncertainties in densities and freeboard. Also, the uncertainty estimates provided with the Envisat SIT data set are possibly too large because the data set producers those days did not adequately take potential correlations between the error contribution into account. Therefore, we rewrote the sentence as: “The large differences in the error bars between Envisat and ICESat mainly result from the inclusion of snow depth uncertainty in Envisat SIT, and lack of adequate regards for potential correlations between the error contribution.”

L208-209: 'ICESat does not capture ... on thicker ice.' I'm not sure where I can see this in Figure 3?

We rewrote the sentence as: “In the eastern Weddell Sea (at sites 229 and 231), ICESat has a few overestimations on thicker ice.”

L210: 'error bars can cover' -> remove 'can'

We removed “can”.

L210: 'However, since many contributions are not well characterized and quantified'. What contributions is this about and how are they not well characterized and quantified?

These uncertainty contributions include spatial and temporal variability on snow depth as well as snow and sea ice density. Few information about these data exists in the Antarctic. Besides, the coverage of sea ice type (first and multi-year) products is incomplete for the Envisat observation period.

L225-226: 'considering the typical sea ice motion'. Briefly characterize this typical sea ice motion (fast, direction?), so the reader can see why the monthly average ULS SIT can be referred to as a spatial average value.

We added the sea ice motion information in the Weddell Sea: The climatological cyclonic atmospheric circulation and the ocean gyre in the Weddell Sea result in westward ice advection along the southern coast and northward advection of ice along the eastern peninsula.

L235: What are 'the ship-based observations'? This is not introduced in the paper before.

The ship-based observations are the ASPeCt data from Worby et al. (2008). According to their Table 3, the average ice thickness in spring West Pacific is 0.68 m, smaller than Envisat and ICESat SIT in our

study.

L237: change 'feature' to 'dissimilarity' or another more descriptive word.

We changed “feature” to “dissimilarity”.

L249: 'but with thickness estimates of up to 1.5 m'. Make sure it is clear to the reader that this is thinner than elsewhere.

We added a sentence to clarify it: “Envisat detects sea ice in the Ross Sea all the years, but with thickness estimates of up to 1.5 m, much larger than expectant seasonal ice thickness.”

L264: 'the two datasets coincide with each other', this sounds a bit like they temporally coincide instead of the distributions being similar (which is I think what you want to say here). Please rewrite.

We rewrote the sentence: “In spring, the two data sets have similar distributions, represented by closer mean and modal thicknesses.”

L269-270: 'We calculate the period-average SIT from the model'. This might be my lack of experience with freezing-degree-days: the FDD in Figure 8 and Table 6 show the total negative temperatures between these months right? I do not understand how it shows SIT. I understand that FDD and SIT are related but I don't see how the model actually calculates average SIT? Please make this more clear in the methods. If 'the model' is not FDD, maybe specify what model you mean?

FDD is calculated by daily degrees below freezing summed over the total number of days the temperature was below freezing. According to Lebedev (1938), a simple model is constructed to produce sea ice thickness:

Thickness (cm) = 1.33 * FDD (°C)$^{0.58}$

Note that the calculated thickness only accounts for the freezing of sea water and excludes ice variations from snowfall, freezing rain or ridging. Therefore, we don’t translate the FDD into thickness growth in this study because i) we think FDD is sufficient to stand for thickness growth and ii) using the very simple translation equation adds uncertainties into our analyses.

L271: 'Envisat SIT has opposite developments from ICESat and FDD during MJ-ON'. Envisat and ICESat do not really show the opposite? They both show the strongest thinning in the western Weddell Sea and both show thickening near the coast in the Amundsen Sea. Please rewrite this to describe the difference, I think something like that Envisat shows more thinning all around the Southern Ocean and ICESat generally more thickening?

We modified the description: “The results show that as FDD is always positive, Envisat SIT shows more thinning all around the Southern Ocean and ICESat SIT generally shows more thickening during MJON, while they share similar variations during FMMJ.”

L271, Figure 8, and Table 6: Please be consistent in how you refer to these periods (MJON or ON-MJ and FM-MJ or MJ-FM). I would suggest for summer to autumn you use FMMJ (instead of the subtraction MJ-FM you used in Figure 8) as this order is more intuitive.

Thanks for your advice. We used FMMJ and MJON to replace the previous usage.

L272: 'both products', which two products? Envisat and ICESat, or satellite altimetry and FDD?

Envisat and ICESat. We changed it to “both satellite products” to clarify.

L274: 'The adverse patterns', adverse (preventing success or development; harmful; unfavourable) might not be the right word here?

We changed “adverse” to “opposite”.

L279-280: 'the regression lines have large positive intercepts in all three seasons, indicating that Envisat SIT tends to exceed ICESat SIT for thin ice'. I can see in Figure 9 that this is true, but the latter does not necessarily follow from the former. A large positive intercept could also be caused by Envisat SIT being

lower than ICESat SIT for thick ice. Again, in the figure I can see this is not the case here, but maybe just rephrase the explanation to just say 'For all five locations, Envisat SIT tends to exceed ICESat SIT for thin ice', without referring to the intercept.

We rewrote the sentence as you suggested: “For all five locations, Envisat SIT tends to exceed ICESat SIT for thin ice.”

L281: change 'splashes' to 'cloud' which is more often used to describe a collection of points in scatterplots. 'Exceed' in what way? Envisat or ICESat or both?

We changed “splashes” to “cloud”. This kind of “exceed” reveals that mean ICESat SIT are nearly constant through all three seasons in the western Weddell Sea, while mean Envisat SIT are noticeably larger in summer and autumn

L284-285: 'The numbers in the last ... values per season'. This might be something to replace to the caption of the table. Also, in the table it does not look like this is in the last column, but in the first row?

We moved this sentence to the caption. We meant to explain the numbers in the last column and we modified the description: “The numbers in the last column of the table are the sample sizes of the comparisons, taking into account the actual number of values per season.”

L294-295: 'it is known that ... homogenous stratigraphy'. This statement could use a citation.

We added a citation: Willatt et al. (2010).

L296: 'considering the large ... of about 70 m'. Maybe specify that the pulse-limited footprint is Envisat and the laser beams ICESat.

We modified the sentence: “considering the Envisat large pulse-limited footprint of about 2–10 km and smaller footprint of ICESat laser beams of about 70 m”.

L341: maybe just say 'may come from the AMSR-E snow depth' here as you haven't yet discussed why it might be biased.

We removed “biased”.

L347-348: 'the differences that AMSR-E snow depths minus the ASPeCt observations are positive ...', rephrase this sentence to something like 'AMSR-E snow depth minus the ASPeCt oservations is positive'

We rephrased the sentence as you suggested.

L349: 'the satellite passive microwave snow depth'. Maybe introduce AMSR-E as a passive microwave sensor in the methods, so readers that don't know the AMSR-E snow depth product know what you are referring to here.

Thanks for your advice. We have introduced AMSR-E snow depth in section 2.1 and 4.2. We just added the passive microwave information in section 2.1: “This snow-depth climatology is derived from the passive microwave sensor Advanced Microwave Scanning Radiometer-EOS (AMSR-E) and AMSR-2 for the Antarctic”.

L351: '... lead to underestimations' and '... lead to overestimations', under- and overestimations of what? SIT?

We meant to say under- and overestimations of snow depth and we added “of snow depth” in the sentence.

L357: The retrieval uncertainty of AMSR-E?

Yes. According to Kern et al. (2015) the average monthly retrieval uncertainty of AMSR-E is 2 cm. However, we realized this value is the precision but not the potential bias which can be much larger. Therefore, we checked sea-ice thickness changes in response to snow biases between 5 and 30 cm in steps of 5 cm.

L357-358: 'suggesting that sea ice thickness change is insensitive to the snow depth', I would suggest change to 'the sensitivity is low', as SIT does change with snow depth, just not by a lot.

We modified the sentence: “suggesting that the sensitivity of sea ice thickness change to the snow depth is low”.

L363-364: 'The usage of snow climatology allows reducing the relative uncertainties', it's a bit unclear what these 'relative uncertainties' are and how they are reduced.

Generally, using a snow climatology for converting ice freeboard to thickness neglects any interannual snow variability. It can reduce the actual snow depth biases to some extent.

L389: Remove 'firstly'

We removed it.

L392: change 'not comparible to' to 'overestimating' or something else more descriptive of the difference between the two.

Thanks for your advice, but we delete this sentence since the previous sentence has stated the result of the comparison with ULS.